WHAT IS CLAIMED IS:

1. A method of comparing nucleic acid sequences being ESTs included in a first database of sequences and nucleic acid sequences included in a second database of sequences to form groups of sequences from the two databases that all relate to the same gene, the method comprising:

for each one or more n-groups of sequences of one of the two databases:

(One) associating therewith lists of nucleic acid sequences, each from one of said two databases, each sequence on the list containing the n-groups; and

(Two) matching sequences on the lists to generate said group.

2. A method for obtaining an mRNA sequence having alternative spliced variants from a database of ESTs, comprising:

providing a raw database comprising a plurality of ESTs; and

assembling ones of said ESTs into mRNA sequences using the method of claim 1, wherein said assembling includes identifying alternative spliced regions.

3. A method according to claim 2, comprising clustering ESTs which have matching segments and wherein said assembly comprising assembling ESTs which are clustered together.

4. A method according to claim 2, comprising correcting errors in said ESTs.

5. An mRNA sequence determined by the process of claim 4.

6. An mRNA sequence according to claim 5, wherein the sequence comprises at least two alternative spliced regions.

7. An mRNA sequence according to claim 5, wherein the sequence comprises at least three alternative spliced regions.

8. An mRNA sequence according to claim 5, wherein the sequence comprises at least four alternative spliced regions.

9. An mRNA sequence according to claim 7, wherein the sequence represents at least two alternative spliced variants of mRNA sequence, each variant utilizing at least one mutually exclusive alternative splice region.

10. An mRNA sequence according to claim 7, wherein the sequence represents at least three alternative spliced variants of mRNA, each variant utilizing at least one mutually exclusive alternative splice region.

11. An mRNA sequence according to claim 7, wherein the sequence represents at least four alternative spliced variants of mRNA, each variant utilizing at least one mutually exclusive alternative splice region.

12. An mRNA sequence according to claim 7, wherein the mRNA sequence is obtained from a single tissue type.

13. A method of mRNA assembly from a plurality of ESTs, comprising:

determining a correspondence between segments in each EST according to the method of claim 1; and

generating a directed graph in which each node represents a single segment, and each transition between two nodes represents the existence of an EST in which the two corresponding segments are consecutive.

14. A method according to claim 13, comprising clustering said ESTs into clusters of associated ESTs, wherein said determining a correspondence is performed on individual clusters of ESTs.

15. A method according to claim 13, comprising identifying alternative spliced regions from said graph based on the morphology of the graph.

16. A method according to claim 13, comprising correcting errors in said ESTs based on said graph based on the morphology of the graph.

17. A method according to claim 16, comprising repeating said clustering responsive to said corrected errors.

18. A method of identifying errors in mRNA sequences, comprising:

generating a graph which represents the assembly of segments of ESTs into an mRNA sequence; and

analyzing said graph to determine unusual configurations of said graph.

19. A method according to claim 18, wherein said analyzing comprises identifying multiple end-nodes in said graph.

20. A method of tuning a database reduction process, comprising:

applying the database reduction process, with a certain value for at least one parameter, to a sample database;

determining a reduction ratio in the database; and

reapplying said method with a new value for said at least one parameter if said reduction ratio is not achieved.

21. A method according to claim 20, wherein said at least one parameter comprises the length of n-groups used in matching two ESTs.

22. A method of EST database processing, comprising:

analyzing said ESTs to detect errors;

further processing said ESTs to create mRNA sequences;

determining, responsive to said further processing, corrections for said errors; and

correcting said errors.

23. A method according to claim 22, wherein said further processing comprises assembling said ESTs into mRNA sequences.

24. A method of designing a DNA chip based on an EST set determined by differential analysis of two biological samples, comprising:

reducing said EST set to a set of mRNA sequences;

analyzing said set of mRNA sequences to determine short mRNA sequences which maximally differentiate said mRNA sequences from mRNA sequences found in both biological samples; and

designing a DNA chip which detects said short mRNA sequences.

25. A method of designing a DNA chip to detect relative expression levels of different variants of mRNA sequences having alternative spliced regions, comprising:

reducing an EST database to determine an mRNA sequence having alternative spliced regions;

enumerating short DNA sequences which are only included in the alternative spliced regions of said different variants; and

designing a DNA chip which detects said short DNA sequences.

26. A DNA chip designed according to the method of claim 24.

27. A method of designing a DNA chip, comprising:

indexing an mRNA database to determine the indexing of short DNA sequences in the mRNA database, which short DNA sequences are of a length suitable for detection by a DNA chip;

determining from said indexing a set of short DNA sequences which uniquely identify a desired mRNA sequence; and

designing a DNA chip which detects said set of short DNA sequences.